

Title	Parallelization of Graph Mining using Backtrack Search Algorithm(Abstract_要旨)
Author(s)	Okuno, Shingo
Citation	Kyoto University (京都大学)
Issue Date	2017-03-23
URL	https://doi.org/10.14989/doctor.k20518
Right	出版社：情報処理学会, IEEE 登録条件：・出典及び利用上の注意事項を明記 ・掲載論文の一部を学位論文に転載する場合は、出典を明記することで可能
Type	Thesis or Dissertation
Textversion	ETD

(続紙 1)

京都大学	博士 (情報学)	氏名	奥野 伸吾
論文題目	Parallelization of Graph Mining using Backtrack Search Algorithm (バックトラック探索アルゴリズムを用いるグラフマイニングの並列化)		
<p>(論文内容の要旨)</p> <p>本論文は、社会ネットワークや遺伝子ネットワークの解析に有用とされる、計算複雑度が高いグラフマイニング問題である、アイテム共有部分グラフ(CIG)抽出問題の並列化手法について論じたものである。本論文では、第1章でCIG抽出問題やその有力な解法であるバックトラック探索とその並列化など、研究の背景について論じるとともに、本論文で示す並列化手法の概要とその計算機科学に対する貢献を概説している。第2章ではCIG抽出問題の効率的な逐次解法であるCOPINEについて論じ、COPINEが実施する枝刈りの論理的な正当性を証明するとともに、その並列化にあたって正当性を担保するための十分条件を示しつつ、基本的な並列アルゴリズムを提示している。また第3章は、本論文で示されるさまざまな実装のために用いたタスク並列言語Tascellについて、その言語機能とそれを実現するための実装方法を概説したものである。</p> <p>続く第4章から第7章までが本論文の中核であり、Tascellを用いたCOPINEの並列実装について、高い並列実行性能を得るために解決すべき技術的な課題とその解決策をさまざまな角度から論じている。まず第4章では、共有メモリ型の並列計算ノードを用いた実装について、効率的な枝刈りを実現するためのタスク間情報共有のメカニズムと、枝刈りを効果的に行うための動的な問題分割・タスク生成戦略を論じ、ワーカ間の負荷均衡と逐次性を有する枝刈りの有効性担保という、相反する性質をもった課題に対する優れた解決策を提示し、その有効性を実証している。また第5章では、分散メモリ型の並列計算環境を対象として、計算ノード間の通信バンド幅や通信遅延による大きなオーバヘッドを考慮した、枝刈り情報の共有メカニズムと複数ノードにまたがる問題分割・タスク生成戦略について論じ、やはり有効な解決策の提示とその実証に成功している。</p> <p>第6章と第7章では、並列探索では不可避である冗長計算をいかに小さく抑えるかという問題を論じている。まず第6章では、冗長な計算を実施しているワーカの処理を中断させるためのメカニズムを、Tascellに例外処理機構を導入することによってエレガントに実現する方法について論じている。また第7章では、この方法を並列COPINEに適用することで、他のワーカによって枝刈りが行われた部分探索木に関する計算を実施しているワーカを発見し、その計算を速やかに中断させることで冗長計算量を小さく抑えるメカニズムについて論じ、性能評価を通じてその具体的な効果を実証している。</p> <p>第8章では、並列バックトラック探索、タスク並列言語、および一般的なグラフ解析・探索に関する関連研究を、本論文の研究と対比しつつ論じている。最後に第9章では、本論文に関する研究成果を概説するとともに、CIG抽出問題などの複雑な並列バックトラック探索をより高速化するために並列言語やその実装が持つべき機能など、今後の研究の展開についても論じている。</p>			

注) 論文内容の要旨と論文審査の結果の要旨は1頁を38字×36行で作成し、合わせて、3,000字を標準とすること。

論文内容の要旨を英語で記入する場合は、400～1,100 wordsで作成し
審査結果の要旨は日本語500～2,000字程度で作成すること。

(論文審査の結果の要旨)

大規模な探索問題の並列化による高速化については、古くから様々な研究が行われており、また並列探索アルゴリズムを簡便に実装するための並列言語の提案も数多く行われている。本論文が対象としているアイテム共有部分グラフ(CIG)抽出問題の並列化と、タスク並列言語Tascellを用いた実装も、これらの並列探索に関する研究の一つとして位置づけられるが、これまでの研究にはない重要な知見をもたらすものとなっている。

まず、本論文で述べている並列探索アルゴリズムのベースである逐次アルゴリズムCOPINEは、探索空間を効果的に限定するための枝刈り操作を、探索木を左・深さ優先で走査することを前提として行っている。その結果、探索の正当性や枝刈りの有効性を保つためには、一般の並列探索では見られない非対称的な枝刈り情報の伝達が求められる。この非対称性を見出し、情報伝達の方法を制限することで探索の正当性が担保されることを証明し、かつ制限付きの伝達を実現する低コストのアルゴリズムとその実装を見出したことは、本論文の重要な貢献と位置付けられる。

次に、枝刈り情報伝達の非対称性が、並列探索における動的な問題分割戦略に強く影響することを見出したことが挙げられる。並列探索で未探索空間を二分して部分問題を生成する際に、何らかの基準で等分とすることが計算負荷均衡の観点で一般的であり、Tascellなどの並列言語もそのような観点で設計されている。しかしCIG抽出問題では、非対称な枝刈り情報伝達を有効に機能させるために、分割比を著しく不均等にすべきことを発見し、これを実現するためにTascellの言語機能を拡張したことは、並列探索アルゴリズムの実装やそのための言語の設計に大きな影響をもたらす成果である。また不均等な分割を行うことによって、これまでは顕在化しなかった自動的負荷分散機構の問題点を明らかにし、それを緩和するための機構を提案してTascellに導入したことも重要な貢献である。さらに、並列探索では一般的に不可避である冗長探索を検知し、それを動的に中断させるメカニズムの提案と、その実装を簡潔に行うために必要な例外処理機能をTascellに付加した成果も、やはり並列探索や言語設計に対して大きな影響力を持つものである。また上記の様々な提案について、現実の遺伝子ネットワークグラフを対象とした評価実験により、それぞれ一定の高速化効果をもたらすものであることを実証している。さらにTascellを含むタスク並列言語に共通する未解決の問題も指摘し、言語の実用性を高めるための指針を提供している。

以上のように本論文は、CIG抽出問題という計算複雑度が高いグラフ探索問題の効果的な並列化を達成しただけではなく、並列探索の方法論やそれを支える並列言語の機能に関して、重要な知見を与えるものとなっている。したがって並列プログラミング一般に関する学術的な貢献度が高く、博士(情報学)の学位に値するものと判定した。また、平成29年2月22日に論文内容とそれに関連した口頭試問を行った結果合格と認めた。

注) 論文審査の結果の要旨の結句には、学位論文の審査についての認定を明記すること。
更に、試問の結果の要旨(例えば「平成 年 月 日論文内容とそれに関連した口頭試問を行った結果合格と認めた。」)を付け加えること。

Webでの即日公開を希望しない場合は、以下に公開可能とする日付を記入すること。
要旨公開可能日： 年 月 日以降